

United States Court of Appeals

FOR THE DISTRICT OF COLUMBIA CIRCUIT

Argued April 20, 1998 Decided August 21, 1998

No. 95-1619

Secretary of Labor,
Petitioner

v.

Keystone Coal Mining Corporation and
Federal Mine Safety and Health Review Commission,
Respondents

Southern Ohio Coal Company, et al.,
Intervenors

On Petition for Review of an Order of the
Federal Mine Safety and Health Review Commission

Edward D. Sieger, Senior Appellate Attorney, United States Department of Labor, argued the cause for petitioner, with whom J. Davitt McAteer, Acting Solicitor, and Nathaniel I. Spiller, Deputy Associate Solicitor, were on the briefs.

R. Henry Moore argued the cause for respondent Keystone Coal Mining Corporation, with whom Heather A. Wyman was on the brief. Norman M. Gleichman, General Counsel, Mine Safety and Health Review Commission, and Elizabeth Ebner, Attorney Advisor, entered appearances.

Timothy M. Biddle, R. Timothy McCrum, H. Thomas Wells, J. Alan Truitt, Laura E. Beverage, L. Anthony George, R. Henry Moore and Heather A. Wyman were on the brief for intervenors Southern Ohio Coal Company, et al. Gail L. Simmons entered an appearance.

Before: Williams*, Sentelle and Randolph, Circuit Judges.

Opinion for the court filed by Circuit Judge Sentelle.
Sentelle, Circuit Judge: The Secretary of Labor ("Secre-

tary"), on behalf of the Mining Safety and Health Administration ("MSHA"), asks us to reverse a November 1996 decision of the Mine Safety and Health Review Commission ("Commission"), affirming rulings by an Administrative Law Judge ("ALJ") in a case involving citations for alleged tampering with coal dust samples. The ALJ and Commission agreed that the Secretary failed to prove (1) in general, that an "abnormal white center" ("AWC") on a coal dust sample filter warrants an inference of intentional tampering; and (2) in a specific test case, that defendant Keystone Coal Mining Corp. ("Keystone") intentionally tampered with its samples. The Secretary argues that the ALJ and Commission held it to an improperly high burden of proof in the first, common-issues proceeding, and that Keystone's exoneration in the second, case-specific proceeding was not supported by substantial evidence. We affirm the Commission's ruling.

I

The case involves over 5000 citations, issued to over 500 coal mines, alleging tampering with air filter samples. These citations arose from a nationwide investigation by the Secretary which began in August 1989. The citations issued between April 4 and June 7, 1991, and included 75 citations to

*Circuit Judge Williams did not participate in this decision. He found it necessary to become recused after hearing oral argument.

Keystone's Urling No. 1 mine ("Urling"). Under the Federal Mine Safety and Health Act of 1977, 30 U.S.C. s 801 et seq. (the "Act"), coal mine operators must periodically sample the concentration of respirable coal dust in the mine atmosphere. The tests employ sampling devices and methods prescribed by the Secretary. The devices are all manufactured by the Mine Safety Appliance Company ("MSA"), and involve basically a pump and a filter cassette. The pump pulls air at a defined rate through the filter, where respirable coal dust is deposited. The filters are then sent to MSHA, within 24 hours of collection. In February 1989, MSHA noticed that some filters had unusual light areas in their centers which generally corresponded to the 6mm opening in the cassette. MSHA concluded that these abnormal white centers were likely caused by reverse air flow--specifically, by a person blowing through the cassette opening in order to dislodge dust from the filter and thereby decrease the sample weight. MSHA expanded the investigation to all mine operators in August 1989, thereafter examining all dust samples for AWCs. Hundreds of mines had no AWCs, but 3900 AWC samples (about 6.5% of all samples received) were discovered by March 19, 1990. On March 20, 1990, MSHA introduced the "AWC void code" which officially notified operators that AWC samples would no longer be accepted as sufficient to fulfill the operator's sampling obligations under the Act. Fewer than 1% of the samples submitted after that date exhibited AWCs.

In August 1992, the ALJ consolidated the citations in order to try common issues (the "common issues" proceeding). The relevant issue in this proceeding was whether deliberate conduct was the "only reasonable explanation" for the cited AWCs. After a 47-day hearing, the ALJ decided against the Secretary, finding that case-by-case inquiry into dust sampling and handling procedures was required to determine whether intentional tampering caused AWCs on samples received from each mine. The ALJ selected Keystone's Urling No. 1 mine for a case-specific trial regarding dust sampling and handling practices. After an 18-day hearing, the ALJ vacated the Urling citations, holding that the Secretary had

failed to prove that Keystone intentionally altered the weight of the 75 cited filters.

The Secretary sought review of both the common issues and Keystone decisions before the Commission. A divided Commission affirmed on November 29, 1995. In re: Contests of Respirable Dust Sample Alteration Citations, Keystone Coal Mining Corp. v. Secretary of Labor, 17 F.M.S.H.R.C. 1819 (1995). Dissenting Commissioner Marks argued that the ALJ had improperly interpreted MSHA regulations to require proof of intentional alteration (an interpretation not challenged here), and further contended that the ALJ had improperly "required the Government to prove that the only cause of the AWCs was intentional conduct, to the exclusion of all other causes!" (emphasis in original). Commissioner Marks would have held that the Secretary had presented sufficient evidence to prevail in both the common-issues and the Keystone proceedings, and that the ALJ's conclusions to the contrary were not supported by substantial evidence. We review the Commission's legal conclusions de novo, *Donovan ex rel. Anderson v. Stafford Constr. Co.*, 732 F.2d 954, 958 (D.C. Cir. 1984), and its findings of fact for substantial evidence, 30 U.S.C. ss 816(a)(1), (b).

II

The Secretary argues that in the common issues proceeding, the Commission and the ALJ erred as a matter of law by requiring a standard of proof higher than a preponderance of the evidence for the proposition that the presence of an AWC allowed an inference of intentional tampering. With respect to the Keystone mine-specific proceeding, the Secretary asserts that the Commission and the ALJ applied an improperly strict burden of proof and that the findings were not supported by substantial evidence.

A

In the common issues proceeding, the Secretary attempted to prove via statistical evidence that the presence of an AWC, without more, established intentional tampering with the

sampling device. Such a finding would have led to a presumption that illegal tampering occurred whenever an AWC was found, perhaps subject to rebuttal by an individual operator who could show that other factors (for example, its handling of filters) caused the AWC in a specific case.

The ALJ held that to prevail the Secretary must prove by a preponderance of the evidence that (1) the AWC definition had a coherent meaning and was consistently applied; (2) the cited AWCs could only result from intentional acts; and (3) the AWCs resulted in weight losses in the cited filters. Although concluding that any inconsistencies in applying the AWC definition were insignificant and that an AWC did result in weight loss, the ALJ found several potential causes of AWCs and received a wide range of expert opinion on the

likelihood of each possibility. For example, AWCs could be caused by tampering, by impact to the cassette, by impact to the air hose, or by snapping together the cassette. The ALJ also found that the likelihood of generating an AWC by non-intentional causes depended upon filter manufacturing characteristics (filter-to-foil distance and filter floppiness), hose pliability, mine and dust characteristics (including type of coal, humidity, weight of dust on the filter, size and shape of particles, and quantity of rock dust or diesel dust on the filter), and cassette population (certain batches of cassettes manufactured by MSA had a greater likelihood of experiencing AWCs, as did all cassettes manufactured before Jan. 1, 1990). Thus, the non-random distribution of AWCs across the mining industry could have been related to tampering at certain mines, but also could have been related to characteristics of certain mine environments or operators' handling techniques.

Therefore, even though the Secretary's statistical evidence demonstrated that AWCs did not occur randomly, the ALJ held that the Secretary had failed to prove that those AWCs were indeed caused by intentional tampering. The Secretary's analysis failed to account for potential accidental causes, manufacturing variables, and mine environment variables. Further, even though the Secretary introduced evidence showing a sharp decline in the number of cited AWCs

in late March, 1990, a date which correlated with the announcement of the "AWC Void Code," the ALJ held that the Secretary had failed to prove that the decline was caused by mine operators responding to that announcement. Thus, the ALJ concluded that the Secretary had "failed to carry [her] burden of proving by a preponderance of the evidence that an AWC on a cited filter establishes that the mine operator intentionally altered the weight of the filter."

The Secretary first contends, as she did before the Commission, that the ALJ imposed an improper burden of proof in this ruling, despite the "preponderance of the evidence" language in both opinions. The Secretary argues that the ALJ erred by requiring proof that "the cited AWCs can only have resulted from intentional acts," Brief for the Secretary of Labor ("Petitioner's Brief") at 41 (emphasis added), or that deliberate conduct "is the only reasonable explanation for the cited AWCs," *id.* (emphasis added). Instead, she contends that she "should have prevailed by establishing on the weight of the evidence that intentional alteration was the more likely explanation for AWCs than other possible explanations." *Id.* (emphasis added). We reject this argument.

In effect, the Secretary sought to establish in the common issues proceeding an evidentiary presumption: that the existence of an AWC, without more, compels (or, at least, allows) an inference that the mine submitting the filter with the AWC intentionally tampered with it in violation of the Mine Act. Such a presumption is only permissible if there is "a sound and rational connection between the proved and inferred facts," and when "proof of one fact renders the existence of another fact so probable that it is sensible and timesaving to assume the truth of [the inferred] fact ... until the adversary disproves it." *Chemical Mfrs. Ass'n v. Department of Transp.*, 105 F.3d 702, 705 (D.C. Cir. 1997) (quoting *NLRB v. Curtin Matheson Scientific, Inc.*, 494 U.S. 775, 788-89 (1990)) (internal citation and quotation marks removed). If there is an alternate explanation for the evidence that is also reasonably likely, then the presumption is irrational.

In making her argument that the evidence presented to the ALJ, and reviewed by the Commission, compelled the imposition of the presumption that every AWC resulted from tampering, the Secretary ignores such cases as *Curtin Matheson and Chemical Mfrs.* She instead relies on *Concrete Pipe & Products of California, Inc. v. Construction Laborers Pension Trust*, 508 U.S. 602, 622 (1993), for the proposition that the preponderance of evidence standard governing the proceedings "simply requires the trier of fact to believe that the existence of a fact is more probable than its nonexistence." It is most evident that the *Concrete Pipe* holding relied on by the Secretary is inapposite. The question before the ALJ, the Commission, and now us, was not whether the Secretary had established by the preponderance of the evidence a simple evidentiary fact--e.g., whether a particular AWC resulted from tampering--but rather whether the Secretary had established that all AWCs result from tampering by some standard sufficiently compelling to require the Commission to adopt it as a presumption. By way of comparison, a plaintiff establishing that a defendant assaulted her is not the same as a litigant convincing a trier of fact that persons similarly situated to the defendant were so likely to have committed assault that liability could be presumed against them.

Unsurprisingly, none of the authorities offered by the Secretary, and none that we have located, hold that a litigant can, even by powerful evidence, compel an adjudicating commission to adopt a presumption favoring the litigant in an entire universe of cases. Generally, the authorities offered by the Secretary and discussed by us concern either the validity or the application of presumptions created either by an administrative body or by statute.

For example, in *Chemical Mfrs.*, we upheld a presumption established by regulation of the Department of Transportation which allowed an inference of inadequate pre-trip inspection from the presence of loose closures on railroad tank cars. 105 F.3d at 703-04. We held that the agency had articulated its reasons for establishing the presumption, and noted that the presumption only shifted the burden of producing evi-

dence. We concluded that the Department had articulated a reasonable evidentiary basis even though it did not consider "every possible intervening event" that could cause a loose closure. Id. at 706. Further, we held that such "administrative presumptions" could be sustained without an evidentiary showing to support the rule, so long as the agency articulates a rational basis. Id. The presumption did no more than "eliminate[] the need to call an expert witness in each enforcement proceeding to establish that properly tightened closures generally do not loosen of their own accord in normal transportation, and that loose closures often reflect inadequate pre-trip inspections." Id. Those facts had been adequately established in the record. We also recognized that because closures were designed "so that, once properly tightened, they will not loosen as a result of vibrations or other conditions normally incident to rail transportation," it was reasonable to presume failure to inspect properly, absent evidence of some intervening event. Id.

The present record does not remotely parallel *Chemical Mfrs.* If an appropriate government agency charged with mine safety regulation had held a rulemaking, established a proper foundation for the presumption advanced by the Secretary, and adopted it, we might well uphold the presumption. At the very least, *Chemical Mfrs.* would be appropriate support for the Secretary's argument. But that is not what happened. A trier of fact took evidence and weighed it. This case turns not on the construction of regulations or on statutory interpretation, but on the weighing of evidence and reasonable inferences made therefrom. Thus, our deference runs not to the policymaking body, MSHA and the Secretary, but to the ALJ, the factfinder who oversees the adjudicatory proceedings.

Curtin Matheson and Concrete Pipe are even less appropriate precedents for this controversy than *Chemical Mfrs.* In *Curtin Matheson*, the Supreme Court reversed the attempt of a circuit court to impose upon an administrative agency the duty of adopting a presumption. 494 U.S. at 781. In no sense did it attempt to set forth terms under which the courts could impose upon an adjudicating commission the

duty to adopt a presumption based upon a certain level of proof offered by a litigant, as the Secretary asks us to do here. Concrete Pipe involved the application of a particular set of presumptions created by statute to a particular sort of factual dispute, 508 U.S. at 630-31, and again offers no support for the Secretary's attempt to impose upon the adjudicators before whom it appeared the duty of presuming.

In another important respect, the presumption sought by the Secretary in this case is far more troubling than the one at issue in Chemical Mfrs. In that case, the Department of Transportation established a rational connection between two concrete facts: the fact of a loose connection allowed inferring the fact that the connection had not been inspected. Absent evidence of an intervening event, such a presumption seems ironclad, especially since a shipper was strictly liable for failure to inspect, without need to prove negligence or intent. But in this case, the Secretary seeks to establish a connection between a fact and an intentional act, namely, to infer from the presence of a light area in a filter's center that the mine operator intentionally and illegally tampered with the sampling device. Distinctions between accidental, negligent, reckless, and intentional conduct, not relevant in Chemical Mfrs., make all the difference between an innocent act and a citable offense in cases involving the Secretary's proposed presumption.

In considering the evidence presented in the common issues proceeding, we cannot say that the ALJ reached an unreasonable conclusion in holding that the Secretary had failed to prove by a preponderance of the evidence that the existence of an AWC established the deliberate conduct required to sustain a citation under the Mine Act and associated regulations. The ALJ certainly did not require that the Secretary prove impossible all other potential causes of AWCs at the hearing. But because AWCs could result from a variety of non-intentional causes, the ALJ found more than a mere "element of doubt" that the Secretary had carried her burden of proof.

To sum up, the Secretary is mistaken in her assertion that under a "preponderance of the evidence" burden, the Commission is required to adopt her presumption when she proves that intentional alteration is merely the "more likely explanation for AWCs than other possible explanations." We therefore affirm the judgment in the common issues proceeding.

B

In the Urling mine-specific proceeding, the Secretary sought to establish by a preponderance of the evidence that Keystone had unlawfully tampered with sampling devices. Both parties introduced a volume of statistical evidence along with the testimony of several experts and witnesses regarding mine conditions and the handling of the filters.

Rochester and Pittsburgh Coal Co. ("R&P") operates 13 mines, including Urling, through several subsidiaries, including Keystone. For all these mines, the independent R&P Environmental Safety Department ("ESD") conducted a coal dust sampling program. From 1970 until 1991, Donald Eget supervised ESD, and Shawn Houck and Douglas Snyder worked with him as laboratory technicians. Normal operating procedures at ESD between 1989 and 1991 had the dust technicians picking up pumps and sampling assemblies in the morning and delivering them to R&P's mines for use that day. Each morning, Eget drove to all 13 R&P mines to retrieve pumps and samples from the previous afternoon and midnight shifts; and each afternoon, the dust technicians returned to ESD with pumps used during the day shift that day. While Eget collected pumps, Houck processed those from the previous day by removing the sampling head and hose, filling out data cards, cleaning the sampled units, recalibrating and reassembling the units, and inserting a new filter cassette. When Eget returned, he inspected the used cassettes, checked the data cards, looked into the inlets and recorded the filter appearances in a logbook for each mine. The cassettes were then packaged and mailed to MSHA.

Robert Thaxton, the MSHA supervisory industrial hygienist responsible for analyzing, monitoring and classifying AWCs, testified that, in his opinion, AWC patterns on Keystone's 75 cited and 3 "no-call" filters resulted from deliberate acts. The Secretary's scientific expert Marple examined and classified the 78 filters, opining that none could result from impact to the cassettes, but that 71 or 72 resulted from reverse air flow, 2 or 3 from a vacuum source introduced to the cassette inlet, and 1 from water introduced into the filter. The Secretary's statistical expert Miller testified that Urling had an AWC citation rate of 43% before the void code notice issued on March 26, 1990 (compared to 6% for other mines), and that the rate dropped to 0.18% after March 26.

Keystone's scientific expert Lee concluded that most of the cited filters indicated lesser forces than would have occurred with deliberate reverse air flow, that the AWC patterns were consistent with a mixed mechanical pulse/reverse air pulse, that humidity reduced the susceptibility to dislodgement, and that water sprays and scrubbers introduced at Urling in 1989 and 1990 contributed to the decline in AWCs. Keystone's statistical expert Roth examined the citation rates of Urling and of all R&P mines combined on a bimonthly basis and concluded (1) that the data showed a gradual decline in AWCs from August 1989 through March 1992, with no significant change in March 1990; (2) that manufacturing variables may have been a factor in AWC formation; and (3) that high incidence rates may be attributable to cassettes manufactured by MSA on four consecutive dates in mid-1989 (for all R&P mines, cassettes manufactured on those four dates were cited at a rate of 50% as opposed to 6% for all other dates of manufacture). Thirty-three R&P employees testified, including ESD personnel Eget, Houck, and Snyder, who described their role in the dust sampling program and uniformly denied tampering or observing anyone else tampering with cassettes.

The Secretary's first argument, much like that advanced and rejected with respect to the common issues proceeding, is that the ALJ improperly held any doubt as to the cause of an AWC sufficient to vacate the citation. Applying such a

burden of persuasion, higher than a "preponderance of the evidence," would constitute reversible error. The Secretary argues that she did prove that tampering was the most likely cause of Keystone's AWCs, even though competing causal theories had not been completely ruled out. In her view, the ALJ should have explicitly determined the probability that rough handling or other non-intentional conduct caused Keystone's AWCs. Without such a determination, according to the Secretary, the ALJ could not have adequately addressed the question of whether the cited filters were more likely than not caused by tampering. We disagree.

The ALJ recognized and the Commission affirmed that the Secretary bore the burden of proving by a preponderance of the evidence that tampering actually occurred, and both agreed that the Secretary had not met that burden. In the process of weighing the vast amount of sometimes conflicting evidence, including the often divergent interpretations by experts, it is simply unreasonable to require that a factfinder determine the mathematical probability of the various different explanations of that evidence. We know of no case in which a reviewing court has required that sort of mathematically nice analysis, nor has the Secretary cited any. Rather, the factfinder must assess whether, on the whole, he is convinced that greater weight of the evidence supports the plaintiff's account. See, e.g., *Steadman v. SEC*, 450 U.S. 91, 101 (1981). So long as that determination is properly made, no further precision or subdivision in specification of probabilities is required. The record indicates such a finding.

The Secretary's second argument reveals the heart of her position: that her evidence showed that tampering was in fact the most likely cause of Keystone's AWCs, despite the ruling of the ALJ and Commission to the contrary. In essence, the Secretary seeks to have this Court review the entire trial record, reweigh the evidence, and decide the case differently. But this Court's duty is to determine whether the findings below were supported by substantial evidence. This sensibly deferential standard of review does not allow us to reverse reasonable findings and conclusions, even if we would have weighed the evidence differently. We must

therefore examine the Secretary's allegations regarding specific inconsistencies between the evidence presented and the conclusions of the factfinder, and determine whether a theoretical "reasonable factfinder" could have reached the conclusions actually reached by the Commission and the ALJ. *United Steelworkers of America v. NLRB*, 983 F.2d 240, 244 (D.C. Cir. 1993).

AWCs Not Random Events

The Secretary presented statistical evidence showing that AWCs were not randomly distributed across all coal mines. Out of samples from 2677 coal mines, about 1300 mines had no AWCs between August 1989 and March 1991. Other mines, like Keystone, had AWCs on more than 40% of their samples submitted during this period. The Secretary insists that this evidence forces the "inescapable conclusion" that "random events do not cause AWCs and AWCs are not inherent in coal mine respirable dust sampling." From this, she concludes that random events (like accidentally dropping a toolbox on an airhose) cannot explain the occurrence of any AWC at any mine, and that the ALJ could not reasonably have relied on random events to explain Urling's high frequency of AWCs.

But the Secretary overstates the record evidence and misunderstands the implications to be drawn from the fact of non-random distribution across mines. Before the ALJ, the Secretary's experts Marple and Thaxton conceded that the Urling AWCs could have been accidentally caused, and that the evidence could not establish whether the pattern on any particular filter resulted from tampering. Miller, the Secretary's statistical expert, did not conclude that intentional misconduct caused the Urling AWCs, but testified only that his conclusions were not inconsistent with tampering.

At best, this evidence demonstrates nothing more than that the likelihood of finding an AWC on a randomly selected filter sample is affected by the mine from which the filter is drawn. In the universe of possible AWC causes, intentional tampering by certain operators is only one of many possibilities that could explain why AWCs occur more frequently at certain mines. Even if all AWCs resulted from purely accidental

causes which were randomly distributed across all mines, the fact that AWC likelihood is affected by environmental conditions like humidity would lead one to expect a non-random distribution of AWCs across mines.

The AWC Rate Decline in Late March, 1990

The Secretary argues that the drop in AWC rates in late March 1990 was statistically significant and interprets it as an indicator of intentional tampering. Because of the correlation between the drop and the date of issue of the AWC void code, the Secretary speculates that Keystone had been tampering but stopped once it learned of the void code. It is undisputed, the Secretary asserts, that Keystone learned of the new void code on March 26, 1990. Miller, the Secretary's statisti-

cian, testified that between August 1989 and March 26, 1990, Keystone's weekly AWC rate fluctuated between 40% and 45%, but after March 26, the rate dropped to near zero and stayed there. According to the Secretary, the "obvious inference" from this is that Eget and Houck, who had sometime earlier learned that MSHA was investigating Keystone's sampling, discovered on that date that MSHA would no longer accept AWCs on dust samples.

The Secretary claims that the ALJ reached his conclusions based solely on Keystone's proffered methodology which analyzed AWC rates on the basis of a bimonthly average. The Secretary argues that there was no good reason for analyzing AWC rates over such a long period, where samples were collected continuously. Of course, such a bimonthly sample interval could make the reduction in AWC rates appear much more gradual, washing out evidence of a sudden change. But the ALJ did not simply adopt Keystone's statistics, as the Secretary argues. Rather, the ALJ weighed all of the statistical evidence, and found that on balance no conclusion could be drawn that there was a dramatic change in AWC rates on or around March 26 that was caused by the issuance of the AWC void code.

On this point, the Secretary advances one reasonable interpretation of the March 26 data. Were we reviewing the evidence de novo, we might (or might not) favor her interpre-

tation. But she falls far short of establishing that the ALJ lacked substantial evidence to reject her interpretation. There is strong evidence that well before March 26, ESD personnel were aware of the MSHA investigation. The ALJ could have reasonably agreed with Keystone that, if truly motivated to stop tampering because of fear of discovery, it would have more naturally done so well before March 26--in fact, 89 citations had already issued by that date. The Secretary responds that Keystone must not have stopped tampering until the date the void code issued, because "[t]he statistical evidence points unequivocally to March 26." This sort of circular argument, assuming the conclusion, is typical of the analysis the Secretary has advanced in this case and does not present an adequate basis to reverse the judgment below.

Even if the March 26 date is ascribed the statistical significance urged by the Secretary, it is a stretch, given the other record evidence, to conclude on that basis that the change in AWC rate is explained by cessation of intentional tampering. The ALJ found that there were other changes around that time, not adequately ruled out by the Secretary's analysis, which also could have lowered the AWC rate. For example, in the relevant period, the ALJ found that there were changes in filter-to-foil distances and other manufacturing variables, increasingly stringent AWC selection criteria, changes in sample handling at Urling, changes in sample handling by ESD personnel, changes in continuous mining machines at Urling, changes in mining conditions, and changes in sampler hose softness. The ALJ evaluated and balanced all these factors to conclude that the Secretary had not demonstrated that any abrupt change occurred on March 26 or that changes in AWC rate justified an inference of prior tampering. The evidence demonstrated that Keystone increased its use of scrubber miners; that the U.S. Attorney's investigation and obvious scrutiny itself might have caused more care in the handling of samples; that R&P heightened its own internal scrutiny and reported actual instances of tampering during this period; that Eget, the roughest handler of pumps at Urling 1, did not transport samples between

April 9 and May 10 because of a bad back; and that after Eget's return, R&P had used up its stock of cassettes with shorter filter-to-foil distances and began using new transport boxes. Thus, substantial evidence supported the ALJ's rejection of the Secretary's interpretation of the declining AWC rate in late March of 1990.

Cassette Manufacture Date

The Secretary rejects as "speculation" the ALJ's conclusion that cassettes manufactured on four consecutive "key dates" in 1989 were responsible for significantly more AWCs. The Secretary contends that when used at mines other than Keystone, cassettes manufactured on those dates actually had a lower than average (2.5%) AWC rate. The Secretary argues that this data suggests nothing more than mere correlation: cassettes manufactured on those dates were used in large numbers when AWCs were occurring at high rates for other reasons. For cassettes manufactured on September 26, 1989 (one of the four dates), 29 of 81 had AWCs before March 26, but 0 of 175 had AWCs after March 26.

The Secretary's expert Miller conceded that the fact that R&P mines had different citation rates with the cassettes from these dates shows only that something is different in the way ESD samples. The Secretary, of course, attributes this difference to intentional tampering by ESD. That is perhaps one reasonable interpretation of the evidence. On the other hand, it is not the only one, and we are obligated not to compel adoption of the Secretary's proffered explanation if the ALJ reached a different conclusion based on substantial evidence.

The ALJ found that Keystone was different from other operators in the way samples were handled and processed. Further, evidence supports the finding that cassettes manufactured on the four key dates in 1989 were responsible for a disproportionate number, over half, of R&P and Urling AWCs. The Secretary's data showed that cassettes from those dates had shorter filter-to-foil distances than later filters, a factor that the ALJ found contributed to the likelihood of a non-intentional AWC. Overall, there is substantial evidence in the record to support the ALJ's conclusion that

the Secretary did not prove by a preponderance of the evidence that intentional tampering, rather than some combination of the filter manufacturing, handling by ESD, and Urling mine characteristics, caused the Keystone AWCs.

Quartz Sampling Data

The Secretary finds error in the ALJ's decision to disregard MSHA data on sampling of quartz between August 1989 and March 1991. Quartz samples were collected in the same fashion and with the same equipment as the coal dust samples, and were transported and processed by ESD in the same fashion. The Secretary claimed that while 44% of the dust samples had AWCs, none of the quartz samples had that appearance. The Secretary's explanation was simple: with

quartz samples, it is not to the operator's advantage to reduce the weight of the dust collected by the device. The ALJ and Commission refused to give any weight to this evidence because the filters at issue were not in evidence, having been destroyed in the normal process of MSHA's quartz analysis; because the filter's appearance had not been preserved through photographs or other records; and because the Secretary had failed to call as witnesses any of the personnel who actually analyzed the quartz filters.

In another attempt to shift the burden of proof, the Secretary notes that one of the actual testers was on Keystone's witness list, but was never called. She forgets that it is the government's burden to prove the existence of a violation. Here, the Secretary introduced no direct evidence--not even photographs or descriptions of the examined filters--to back up these claims. The ALJ and the Commission did not err in refusing to draw any conclusion from this evidence.

ESD Employee Conduct and Witness Testimony

The Secretary introduced direct testimony regarding ESD employees looking into dust filters and talking about what might happen if they blew into them, arguing that this evidence justified the conclusion that they were in fact blowing into them. Keystone offered the testimony of the employees who handled the cassettes to the effect that they did not tamper with them. The Secretary asserts that the ALJ

should not have credited ESD employees' denials of tampering. The Secretary describes as "insupportable" the ALJ's stated reasons for crediting Eget and Houck: the absence of motive for tampering and the strong disincentive from their knowledge of possible sanctions. The Secretary also argues that these witnesses contradicted themselves and each other. The Secretary asserts that it was error for the ALJ to have believed those witnesses, arguing that "[b]ecause denials of tampering by ESD witnesses are inconsistent with the other evidence, the ALJ's credibility findings would not stand even if they had been based on demeanor." For this remarkable proposition, the Secretary cites two cases, *Bishopp v. District of Columbia*, 788 F.2d 781, 785-86 (D.C. Cir. 1986); and *Millar v. FCC*, 707 F.2d 1530, 1539 (D.C. Cir. 1983). Unsurprisingly, neither of these cases goes anywhere nearly so far as to say that a trier of fact commits error by believing a witness whose evidence is inconsistent with other evidence. Logically, of course, the Secretary's proposition could not stand. If evidence could not be credited when it was contradictory to other evidence, then presumably neither could the other evidence be credited since it is contradictory to that rejected in the first instance. As one might expect, neither the *Bishopp* nor the *Millar* case stands for the proposition which the Secretary asserts.

What we actually held in *Bishopp* was that "we must be particularly careful to defer to the district court's credibility findings...." 788 F.2d at 786. Obviously, that is the very opposite of what the Secretary asserts. With due charity to

the Secretary, we note that we went on to say that in "the rare case" we would reverse even "under this very restricted scope of review," when "the judge below credited a witness whose testimony was so internally inconsistent or implausible on its face that a reasonable factfinder could not credit it." Id. Millar is to the same effect, allowing for a reversal where a witness's testimony is "so incredible," or is faced by "contrary evidence ... so overwhelming," that a reasonable factfinder could not believe the testimony regardless of the witness's demeanor. 707 F.2d at 1539. Thus, both of the cases upon which the Secretary relies are little more than

restatements of the "reasonable factfinder" standard of review as applied to credibility determinations. We have alluded to the Secretary's misunderstanding of that standard above, and will discuss the same further infra. As to this argument, it is sufficient to say that the Secretary has fallen far short of that demanding standard.

The record demonstrates that the ALJ specifically and carefully assessed the credibility of the employee witnesses, and found that their denials of tampering were not only believable, but consistent with other evidence. The Secretary simply has not explained to this Court why we must depart from the rule that a factfinder's determinations of credibility are entitled to great deference. Nothing justifies the extraordinary step of overturning these findings. See *Chen v. GAO*, 821 F.2d 732, 738 (D.C. Cir. 1987).

Accidents, Rough Handling, Filter Manufacturing, Mine Environment

The Secretary would have us reject the ALJ's findings that accidents and rough handling of samples could have contributed to or explained Keystone's high AWC rate. These conclusions were based on the testimony and theories of Keystone's scientific expert Lee, who concluded that handling could have accounted for many AWCs; that short filter-to-foil distance increased AWC likelihood; that increased humidity decreased AWC likelihood; that increased use of scrubbers made it more difficult to dislodge dust from filters, decreasing the AWC rate; and that the AWCs on Urling filters resembled dislodgements caused by impact, not reverse air flow. The Secretary argues every detail of the evidence at length. In essence, she contends that her scientific evidence was so overwhelmingly correct and so clearly compelled her conclusion that the ALJ could not lawfully have found against her. But the record does not support this proposition.

The ALJ found that in many instances the Secretary's scientific evidence was inconclusive or otherwise could not be adequately evaluated. All of these issues involve conflicting expert testimony, and this Court must defer to the reasonable determination of the trier of fact regarding not only the relevance but the reliability of expert testimony presented at

trial. See *General Electric Co. v. Joiner*, 118 S. Ct. 512 (1997); *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 509 U.S. 579, 589 (1993); see also *Millar*, 707 F.2d at 1539. We cannot deem unreasonable the conclusion that the Secretary failed to meet her burden of proof.

The record clearly supports the proposition that accidental events caused at least some citable AWCs at Urling. Hose impacts, for example, occurred routinely in the transport of sampling apparatus. Further, because the cassettes were not removed and transported separately from the testing apparatus, R&P's filters may have had a greater potential for such impacts than other mines. And, the Secretary conceded that filter samples collected by MSHA field personnel sometimes contained AWCs, apparently caused by opening and reclosing

of the filter cassettes at MSHA. Although there was apparently no evidence on this point, it is at least possible that the Keystone filters might have been opened and reclosed after delivery to MSHA. The likelihood of these various possible causes cannot be established with mathematical precision. The Secretary's burden is to demonstrate, by a preponderance of the evidence, that intentional tampering actually caused the dust dislodgment on the particular filters at issue in each citation. The ALJ and Commission reasonably concluded that she had not carried that burden.

III

Ultimately, the Secretary's position is fraught with misunderstanding of the nature of her burden of proof and of the danger of relying on a probabilistic estimate of the correlation between some observation and a proffered explanation of its cause. In the first instance, the Secretary never seems to accept the fact that we review this case under the standard of the reasonable factfinder. That standard, as we have noted, renders the Commission's "findings of fact ... 'conclusive' when supported by substantial evidence on the record considered as a whole." *United Steelworkers*, 983 F.2d at 244. Occasionally, though rarely, we do hold that record evidence is not sufficient to support a decision in favor of a party with

the burden of proof, even in the face of that deferential standard. Even less frequently have we held that evidentiary support for the party with the burden of proof was so overwhelming that a trier of fact erred by ruling that the burdened party had not carried its load. The Secretary has pointed to no such case and our research has uncovered only one. See *Gibson Greetings v. NLRB*, 53 F.3d 385, 389 (D.C. Cir. 1995).

The closest the Secretary comes is *Bishopp*, supra. There, the plaintiff had lost in the trial court in an employment discrimination case. Obviously, the plaintiff ultimately bore the burden of persuasion. *St. Mary's Honor Ctr. v. Hicks*, 509 U.S. 502, 507-11 (1993). But in *Bishopp*, the plaintiffs had presented a prima facie case under *McDonnell Douglas Corp. v. Green*, 411 U.S. 792 (1973). The issue upon which we reversed the district court was whether the defendant had come forward with legitimate nondiscriminatory reasons for the commission of the allegedly discriminatory acts which made out the prima facie case. On that issue, the appellees had borne at least the burden of production, and it was on that issue that we reversed. *Bishopp*, 788 F.2d at 789. This is not to say that we would never find a record so overwhelming as to require us to "direct a verdict" in favor of the party with the burden of proof, but it is to say that given the deferential standard of review such a case would be rare indeed. This is not such a case.

Although she picks at various items of evidence, the Secretary principally relies on her evidence of probability--that it was more likely than not that the cause of any given AWC was intentional tampering. This falls far short of the compelling case in which a reasonable finder of fact must find for the party with the burden of proof in the face of direct evidence supporting the other litigant. There is a false sense of security that comes from the use of numbers, which in this context can appear much like scientific data. But any useful scientific measurement must be accompanied by an estimate of its uncertainty, and when the entire body of evidence has been considered, the Secretary fails to persuade that she has established with any certainty that AWCs in general, or

Keystone's AWCs in particular, were in fact caused by intentional tampering.

Over and over, the Secretary insists that she established that the mathematical probability of tampering was something greater than 50%. Arguing from precedents involving employment discrimination, she contends that similar statistical evidence may be deemed sufficient to establish a prima facie case of intentional discrimination or to rebut a defendant's explanation as pretextual. See *Palmer v. Shultz*, 815 F.2d 84, 90 (D.C. Cir. 1987); *McDonnell Douglas*, 411 U.S. at 792. Statistics alone may suffice to show illegal discrimination "if they are condemning enough," *Berger v. Iron Workers Reinforced Rodmen Local 201*, 843 F.2d 1395, 1413 (D.C. Cir. 1988) (citation omitted), and cannot be dismissed "on mere conjecture," *Palmer*, 815 F.2d at 106. The Secretary notes that in those cases, a result more than two standard deviations from the mean (indicating a 95% probability that the relationship is not random) suffices in most instances to give rise to an inference of intentional action. *Berger*, 843 F.2d at 1412.

These precedents lend little aid to the Secretary's cause. Statistics may show a correlation between some characteristic (for example, age) and some unequal treatment (for example, refusal to hire), yet a finding of discrimination is allowed only (1) if the employer fails to present a legitimate justification or (2) if the factfinder concludes that the greater weight of the evidence, including the statistical data, supports a conclusion that the particular employee suffered illegal discrimination. In situations where direct evidence is difficult or impossible to obtain, a party may meet his burden of proof with statistical evidence alone. (This may account for its acceptance as such in some employment discrimination cases. See, e.g., *Berger v. Iron Workers Reinforced Rodmen Local 201*, 843 F.2d at 1413; *Palmer v. Shultz*, 815 F.2d at 90.) Even then, statistics must reasonably control for a variety of factors to properly define similarly situated employees, and in any event may be counterbalanced by evidence providing an alternate explanation of the pattern or of the particular action in question. The weight given to statistical evidence in such cases is not

absolute, but depends on the degree to which it rules out legitimate explanations and how the statistics factor into the balance with the other available evidence. See, e.g., *Coward v. ADT Security Systems*, 140 F.3d 271, 276-77 (D.C. Cir. 1998) (Sentelle, J., concurring). Here, it is true that AWCs are not randomly distributed across all mines, and that something probably explains the higher frequency of AWCs at Urling. But without direct evidence of tampering, and given the substantial basis in the record for alternate theories, there are no statistics "condemning enough" to require reversal of the judgment below.

The Secretary throughout this case assumes that proving probability is the same thing as convincing a trier of fact by the greater weight of the evidence. While the two propositions may sound superficially similar, they are not the same. This case well illustrates why. When the Secretary has cited a responding mine for tampering with a particular filter, certainly evidence of the probability of the cause of the AWC on that filter is relevant. This relevant evidence does not mean that the trier of fact must be convinced to any degree that the mine operator's employees tampered with that particular filter. An hypothetical that reverses the facts of this case demonstrates why. If it were the burden of the mine operators to prove their innocence, and they came forward with evidence that 99% of all filters had never been tampered with, this would not mean that they would be entitled to an acquittal as to particular filters on which the Secretary could offer direct evidence of tampering. For example, if the same witnesses who came forward here to testify that they had committed no such acts instead came forward and swore that "we tampered with these filters," we could hardly say that a reasonable trier of fact would have to disbelieve them because statistical data proved that such tampering was extremely unlikely. The same is true here.

Perhaps the Secretary is right that a majority of the AWCs were caused by tampering. Perhaps she is not. Either way, it is not unreasonable for the finder of fact to conclude that the Secretary did not establish that a particular filter in evidence fell into the majority rather than the minority group.

To offer one further hypothetical illustrative of the Secretary's misconception, we recall the example created by Professor L. Jonathan Cohen. He posits a situation in which uncontroverted evidence establishes that something over half of 1000 attendees at a rodeo entered without paying the admission fee. He rightly concludes that even though that evidence suggests that it would be "more likely than not" the case that a randomly selected attendee had not paid, that evidence would be legally insufficient to allow judgment against a specific selected attendee for the price of admission. Most likely such evidence without more would not even be submitted to a jury. See L. Jonathan Cohen, *The Probable and the Provable* 75 (1977). In our case, the problem is not merely that it is difficult to state with precision the probability that a randomly selected AWC was caused by intentional tampering. The problem here, as in the gatecrasher hypothetical, is that the uncertainty arising from all of the information not presented to the factfinder (e.g., evidence regarding potential alternative causes for each AWC, its course of handling, mine conditions, and so forth) is of such degree that the factfinder cannot confidently say that the weight of the evidence supports the proposition. In other words, the weight ascribed to the evidence is affected, in part, by the factfinder's judgment about the volume and significance of relevant information that is not available for examination. See Neil B. Cohen, *Conceptualizing Proof and Calculating Probabilities: A Response to Professor Kaye*, 73 *Cornell L. Rev.* 78, 86 (1987) ("Convincing the factfinder of such a probabilistic judgment requires more ... than simply noting that the best guess of the probability exceeds 0.5; rather, ... the factfinder also takes into account its judgment as to how likely the best guess is to 'hold up.' ").

Conclusion

In each of these proceedings, whether we would have reached the same conclusion as the ALJ is irrelevant. We might have upheld a ruling in favor of the Secretary on the basis of this record. But the Secretary has not come close to proving that the decisions below were unreasonable or not

supported by substantial evidence. Indeed, we find it highly unlikely that the government would desire a standard of review that would allow us to reverse such a decision based on nothing more than our distant and inexperienced view of the record evidence. We therefore affirm the decision of the Commission and deny the petition for review.

United States Court of Appeals
For The District of Columbia Circuit

No. 95-1619

September Term, 1997

Secretary of Labor,
Petitioner

v.

Keystone Coal Mining Corporation and Federal Mine
Safety and Health Review Commission,
Respondents

Southern Ohio Coal Company, et al.,
Intervenors

O R D E R

It is ORDERED, sua sponte, that the opinion filed herein this date is amended,
as follows:

On Page 2, in the listing of judges

insert * next to the name Williams

Add footnote to page 2, as follows

*Circuit Judge Williams did not participate in this decision. He found it necessary to become recused after hearing oral argument.

FOR THE COURT:
Mark J. Langer, Clerk
BY:
Robert A. Bonner
Deputy Clerk

Filed August 21, 1998